

Juin 2012

ATELIER EPIGENETIQUE

Utilisation de la Q-PCR pour analyser des données de CHIP ou de MeDIP

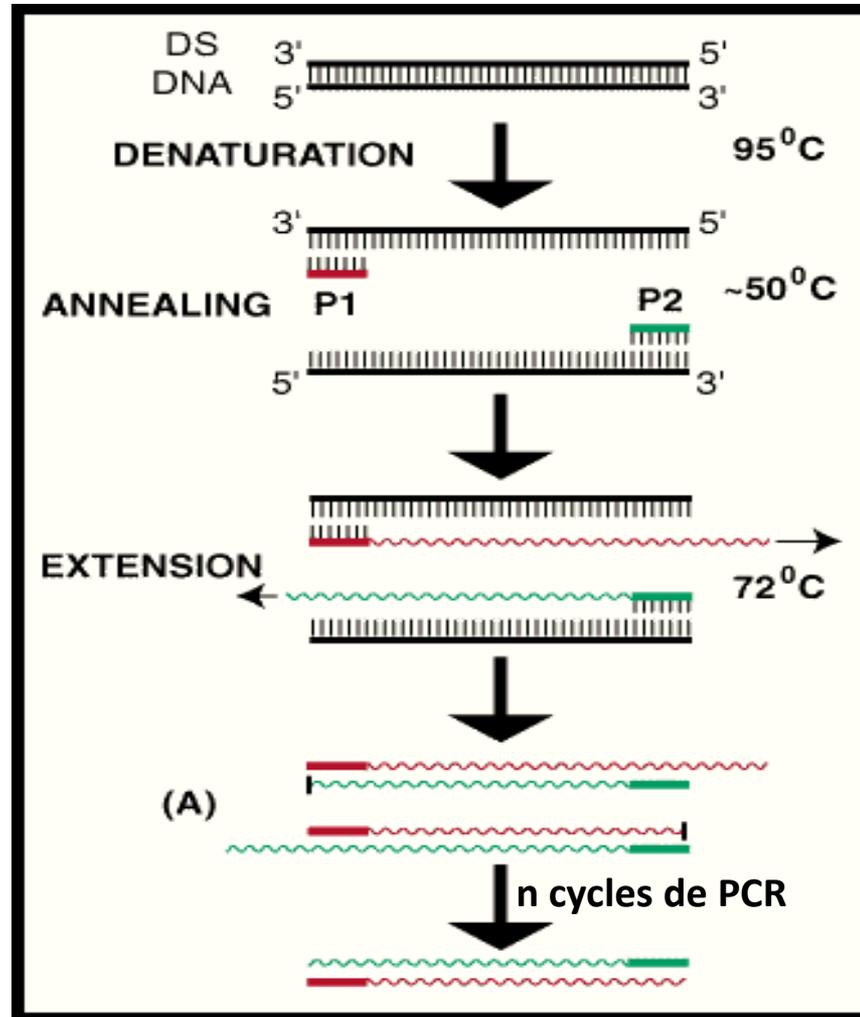
Emmanuèle Mouchel-Vielh

I. La PCR quantitative: principe et généralités

II. Interprétation des résultats de Q-PCR pour les expériences de CHIP et de MeDIP

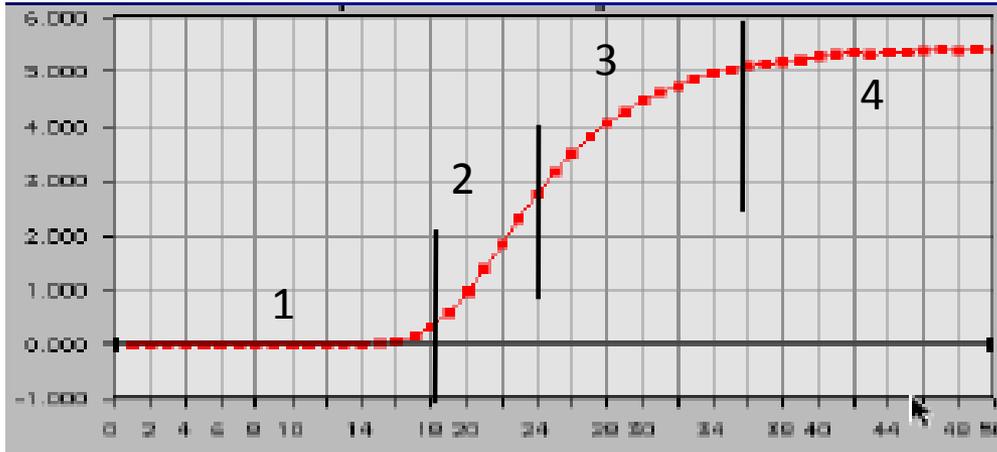
La technique de PCR

X_0 copies de la séquence cible au départ



X_n copies de la séquence cible à la fin

Equation de la PCR, courbe primaire, C_T



C_T = valeur du cycle à partir duquel le produit devient détectable

Equation de la PCR:

$$X_n = X_0 AE^n$$

AE: efficacité d'amplification $0 < AE < 2$;
n: nombre de cycles

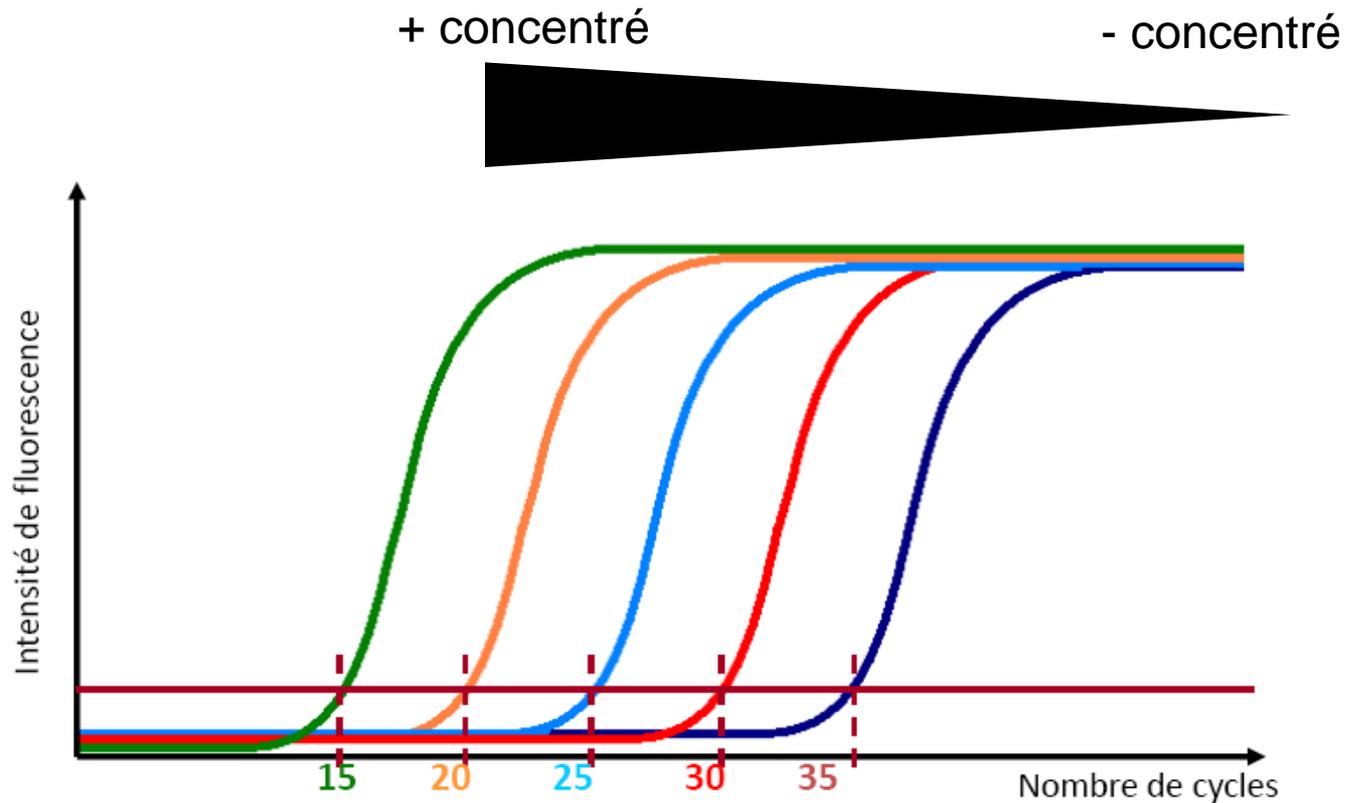
Si $AE=2$, $X_n = X_0 2^n$

- ✓ Phase 1: amplification exponentielle mais produit non détectable
- ✓ Phase 2: amplification exponentielle
- ✓ Phase 3: amplification, phase « log-linéaire »
- ✓ Phase 4: phase de plateau, plus d'amplification (saturation)

➤ **PCR classique (en point final):** on regarde les produits au plateau (non quantitatif)

➤ **PCR quantitative:** on quantifie les produits pendant tout le déroulement de la PCR (quantitatif dans la phase exponentielle)

Valeur de C_T et quantité d'ADN initiale



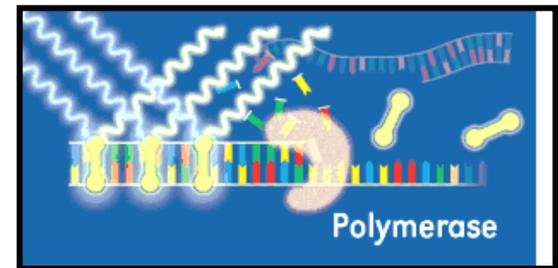
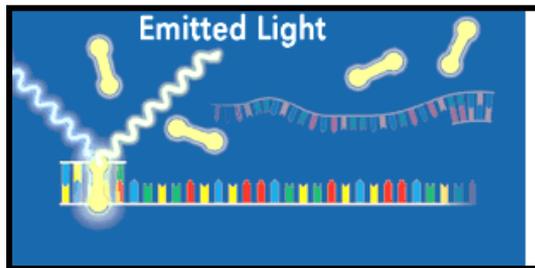
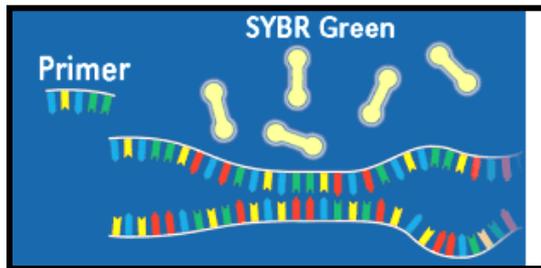
- Plus la quantité d'ADN initiale est élevée, plus la valeur de C_T est faible
- Pour $AE=2$, une différence de 1 C_T correspond à un facteur 2 en quantité d'ADN

Propriétés générales de la PCR quantitative (= en temps réel)

- ✓ Détection à partir de 100 copies d'une séquence dans l'échantillon (environ 10^{-15} g)
- ✓ Quantification absolue ou relative
- ✓ Très nombreuses applications (génotypage, étude de l'expression des gènes, chIP...)
- ✓ Pour avoir la meilleure efficacité possible: amplicons de petite taille (70 à 250/300 bp)
- ✓ Principe général de détection: utilisation d'un (ou plusieurs) fluorochromes: l'émission de fluorescence est le reflet de la quantité d'ADN synthétisée à chaque cycle
- ✓ Principe des appareils de Q-PCR: un thermocycleur couplé à un fluorimètre qui mesure la fluorescence émise à chaque cycle (fluorimètre= laser pour exciter les fluorochromes+ système de détection de la fluorescence émise)
- ✓ Etablissement de la courbe primaire de PCR : émission de fluorescence=f(n). Cette courbe permet de déterminer la valeur de C_T , qui permet ensuite d'estimer (de manière absolue ou relative) la quantité X_0 présente au départ dans l'échantillon.

Q-PCR en SYBR-Green

- ✓ **SYBR Green**: intercalant de l'ADN: 20 fois plus fluorescent quand il est intercalé que sous forme libre
- ✓ A chaque cycle, en fin de phase d'élongation: excitation/mesure de la fluorescence émise



Avantages et inconvénients de la Q-PCR en SYBR-Green

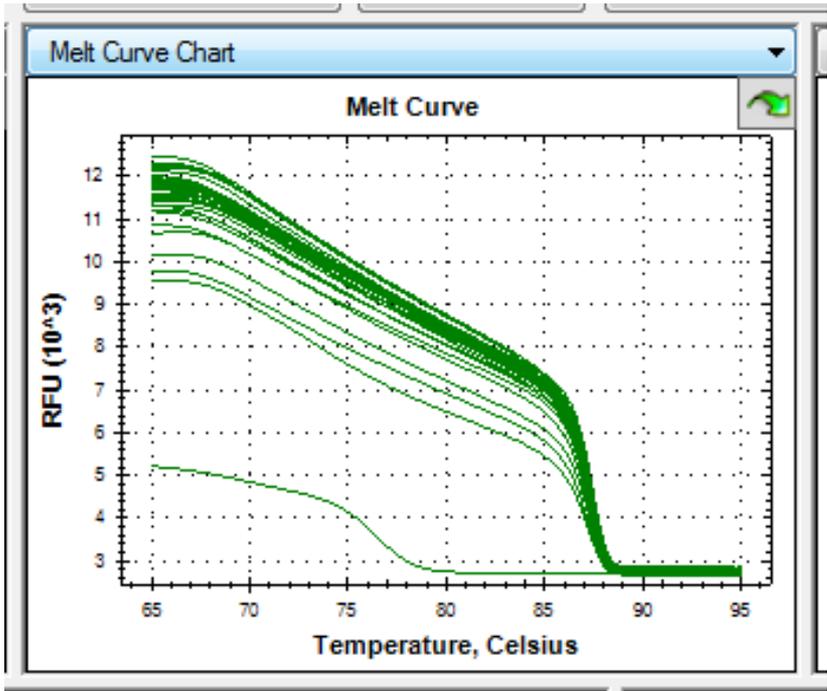
Avantages:

- ✓ Méthode simple à mettre en œuvre, en particulier au niveau du choix des amorces
- ✓ Méthode la moins chère

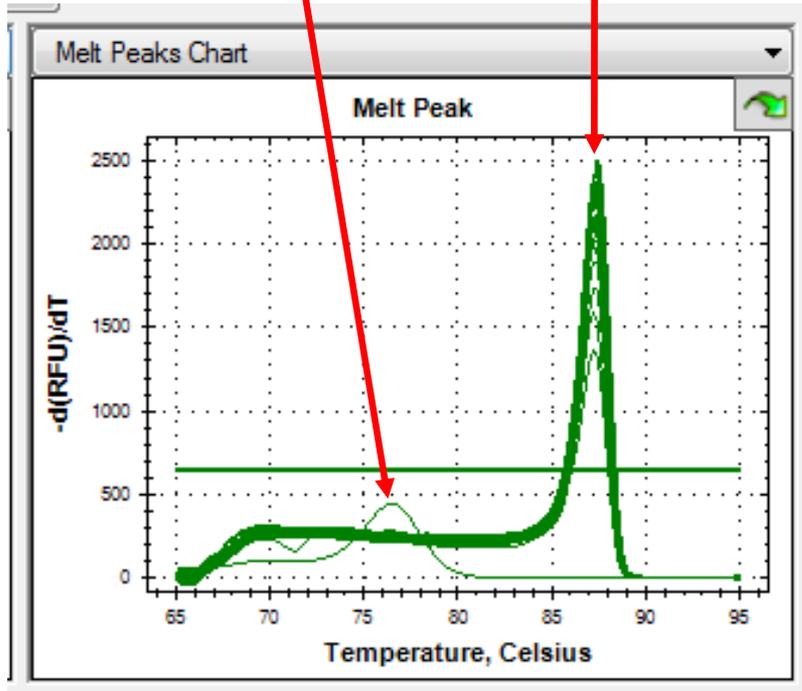
Inconvénients:

- ✓ Ne permet pas de faire des PCR multiplex (=amplifications de plusieurs séquences dans le même tube)
- ✓ Moins sensible que d'autres méthodes (détection à partir de 1000 copies)
- ✓ Risque possible d'aspécificité (production d'amplicons aspécifiques, quantifiés comme les amplicons spécifiques): pour vérifier, réalisation d'une courbe de fusion en fin de PCR:
1 pic= 1 amplicon (possibilité d'observer les dimères d'amorces)

Courbe de fusion



Fluorescence=f(T)

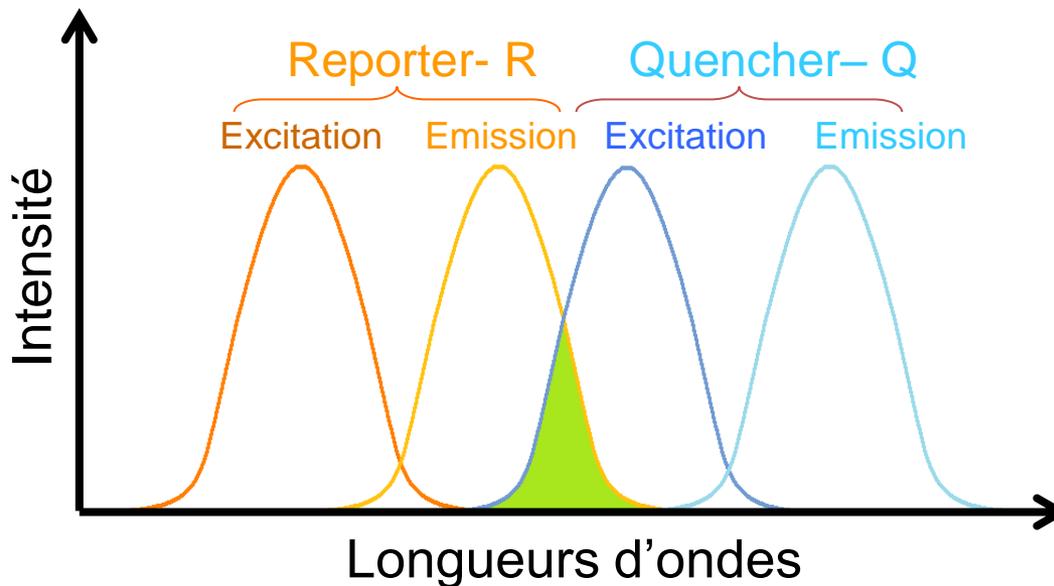


d''(Fluorescence)=f(T)

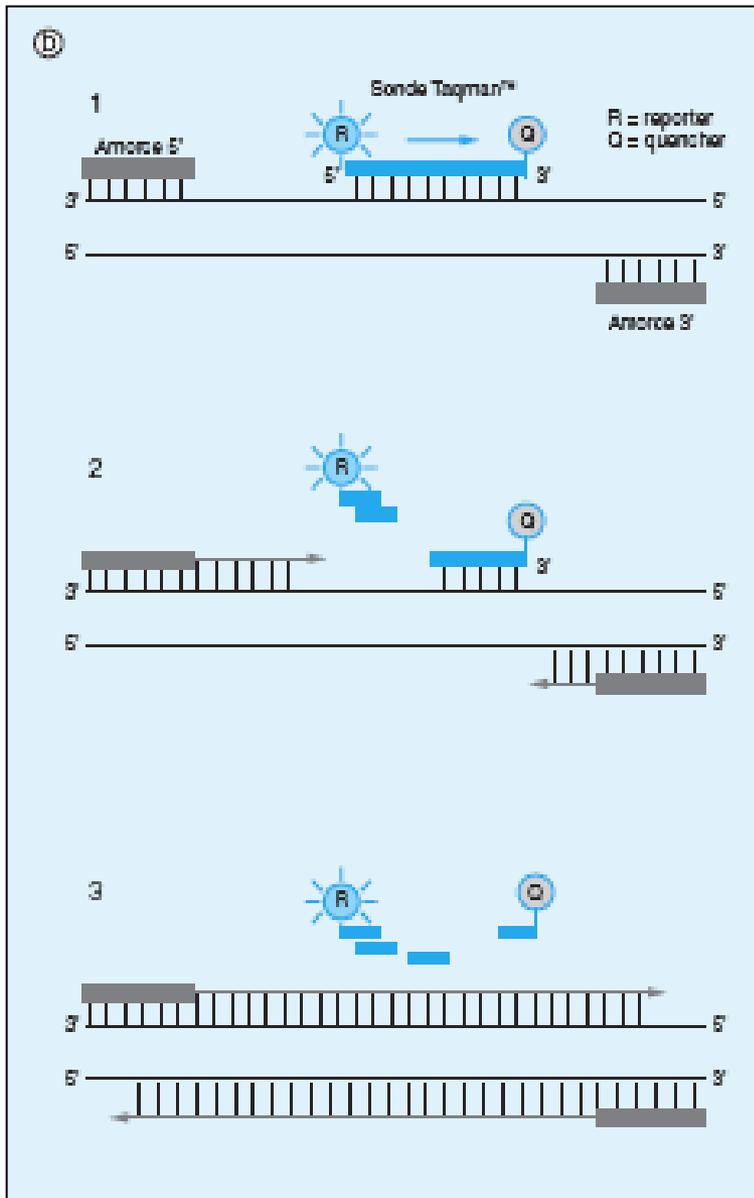
Q-PCR et technologie FRET: utilisation de sondes TaqMan

FRET= Fluorescence Resonance Energy Transfer

- ✓ utilisation de deux fluorochromes à spectre chevauchant: le spectre d'émission du Reporter recouvre le spectre d'excitation du Quencher
- ✓ quand R et Q sont proches: Q capte la fluorescence émise par R



- ✓ Sondes TaqMan: principe du FRET et utilisation de l'activité exonucléase 5'-3' de la Taq polymérase (= sonde d'hydrolyse)



✓ Utilisation de 2 amorces et d'une sonde TaqMan: 20 à 40nt, interne aux 2 amorces, spécifique de la séquence à amplifier, et marquée en 5' par un fluorochrome R et en 3' par un fluorochrome Q; T_m plus élevé que celui des amorces

✓ En cours d'élongation, la sonde est lysée: séparation de R et Q

✓ A la fin de l'étape d'élongation, excitation de R et mesure à λ émission R: pour chaque molécule synthétisée, un R émet de la fluorescence

Avantages et inconvénients de la Q-PCR en sondes TaqMan

Avantages:

- ✓ permet les PCR multiplex (utilisation de différents couples Q-R)
- ✓ élimine le problème des hybridations aspécifiques
- ✓ plus sensible que la méthode SYBR-Green

Inconvénients:

- ✓ Moins facile à mettre en œuvre que la méthode SYBR-Green (qualité de la sonde TaqMan)
- ✓ Plus cher que la méthode SYBR-Green

Exemples d'applications de la Q-PCR

- ✓ Analyse du niveau d'expression des gènes, mise en évidence de variants d'épissage... (QRT-PCR)
- ✓ Détection et quantifications d'agents pathogènes
- ✓ Génotypage, étude de polymorphismes
- ✓ **Analyse d'expériences de CHIP ou de MeDIP**
- ✓

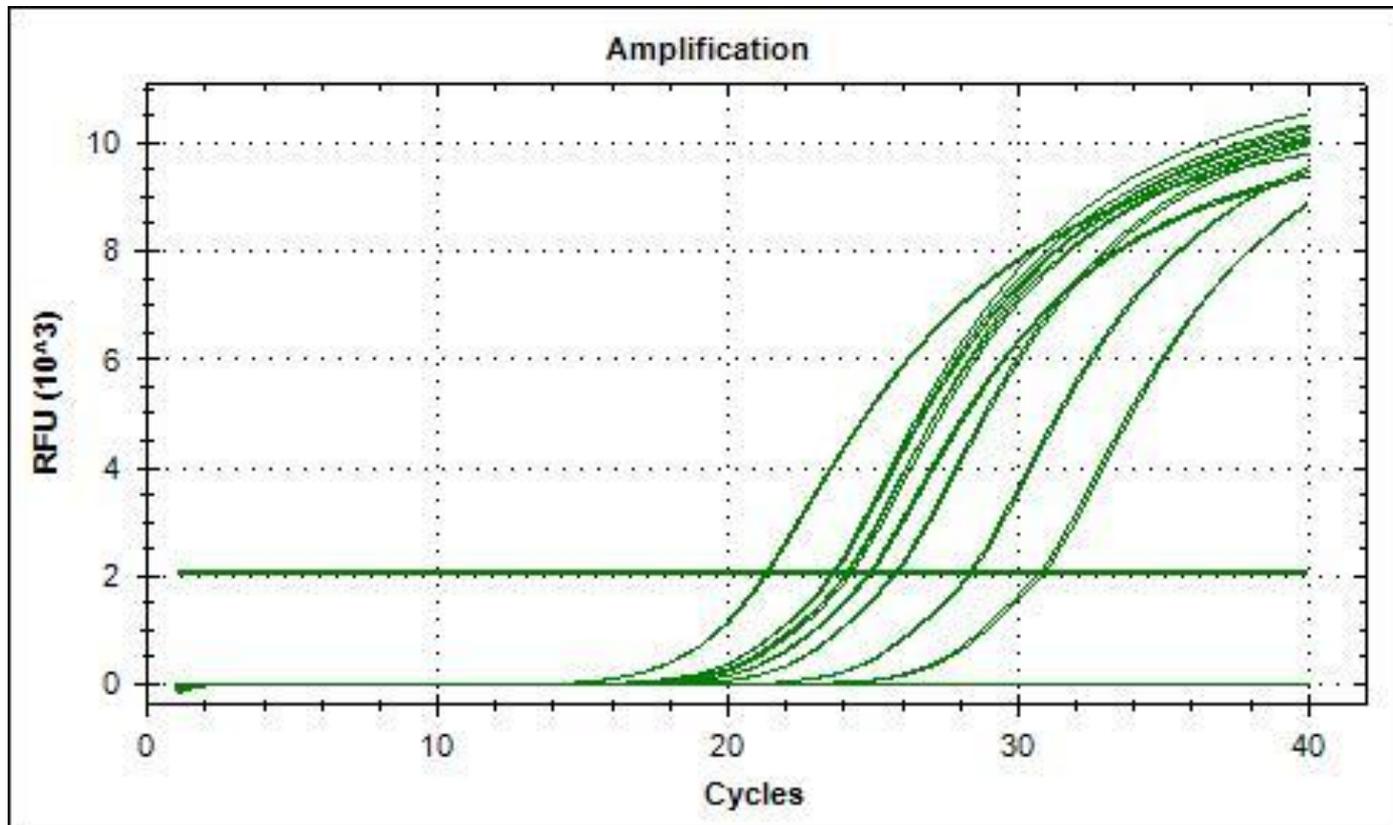
Interprétation des résultats de CHIP et MeDIP par Q-PCR

- ✓ Choix du système de fluorescence (de préférence TaqMan, mais possible aussi en SYBR-Green)
- ✓ Choix de la séquence à amplifier: petite taille des amplicons (50 à 150 bp)
- ✓ Choix des amorces: critères classiques (absence de structure secondaire, pas d'appariement entre amorces, spécificité)
- ✓ Validation des amorces: réalisation d'une **courbe de calibration** pour vérifier leur efficacité (calcul de AE): réactions de Q-PCR (en triplicats) sur des dilutions en série (1/5 ou 1/10) d'un ADN contenant la séquence matrice à amplifier (ADN génomique de préférence: non nécessaire de connaître la concentration de la séquence matrice dans cet ADN): **AE doit être proche de 2**

Evaluation de l'efficacité de la PCR et détermination des valeurs X_0 dans les échantillons

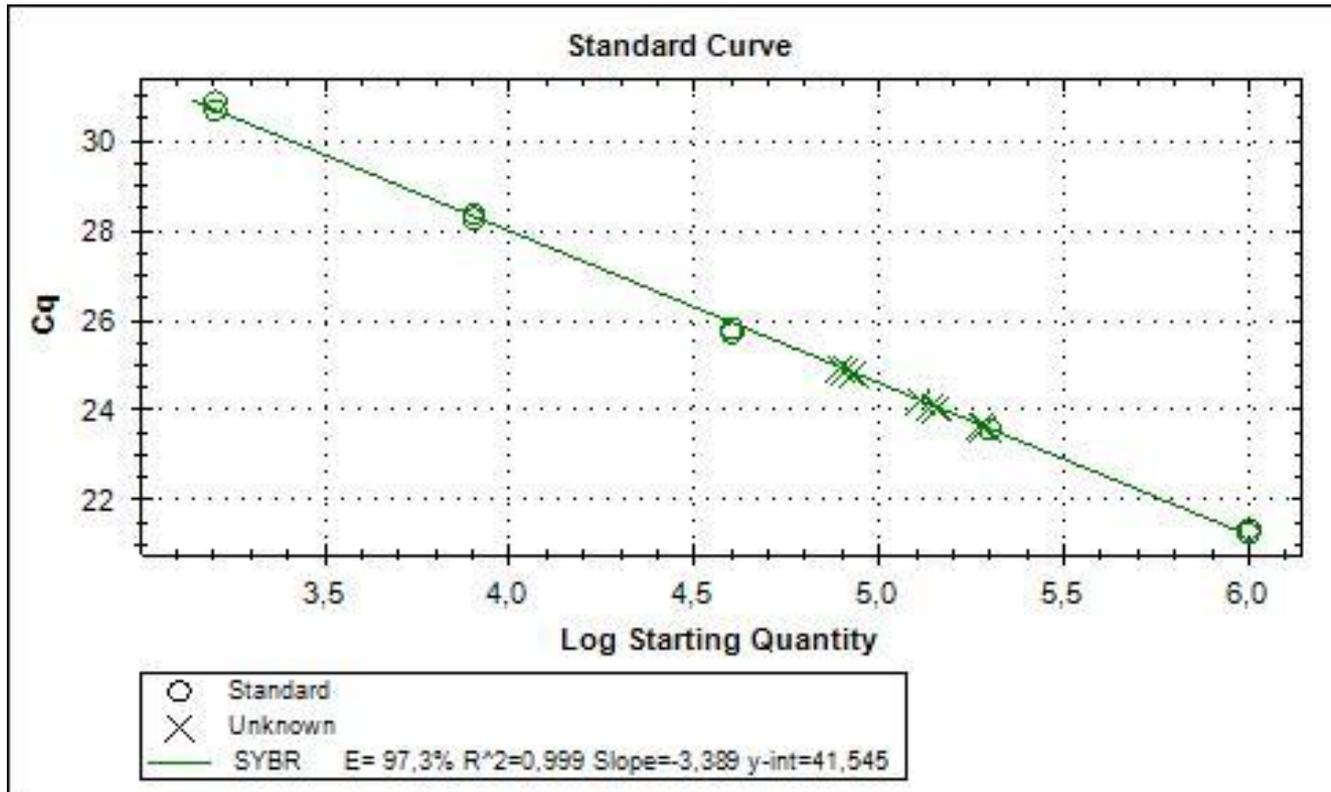
Courbe primaire: Fluorescence=f(nb de cycles)

Permet de déterminer la valeur de C_T pour chaque point



Courbe standard: $C_T=f(\text{Log } X_0)$

- Etablie à partir des points de la gamme de calibration.
- Permet de déterminer le coefficient de corrélation r^2 (doit être supérieur à 0,985)
- Permet de déterminer l'efficacité de la réaction (doit être comprise entre 90% et 105%)
- Permet de déterminer la valeur de X_0 dans les échantillons, à partir de leur valeur de C_T



Equation de la réaction de PCR:

AE= Amplification Efficiency $0 < AE < 2$

$$X_n = X_0 * AE^n$$

Pour $n=C_T$:

$$X_{C_T} = X_0 * AE^{C_T}$$

$$C_T = \frac{\text{Log } X_{C_T} - \text{Log } X_0}{\text{Log } AE}$$

$$C_T = \text{Log } X_0 (-1/\text{Log } AE) + \text{Log } X_{C_T} (1/\text{Log } AE)$$

$$C_T = a \text{Log } X_0 + b \quad \text{Pente } a = -1/\text{Log } AE$$

$$AE = 10^{(-1/a)}$$

- Pour $a = -3,32$ $AE = 2$
si $AE = 2$, efficacité = 100%

Calculs pour les expériences de ChIP et MeDIP (1)

- Faire la moyenne des CT des duplicats ou triplicats
- Calcul du pourcentage d'input immunoprécipité par l'anticorps: comparaison des CT, en corrigeant le CT de l'input par rapport au facteur de dilution (FD= facteur de dilution de l'input):

- ChIP: dilution 1: 15/100 (p23); dilution 2: 20/100 (p33); total: 300/10000;
 $FD=10000/300=33,3$

- MeDIP: dilution 7,5/100 (p29); $FD=100/7,5=13,3$

$$\% \text{ input} = AE^{(CT_{\text{input_corrige}} - CT_{\text{IP}})} \times 100\% = AE^{\Delta CT} \times 100$$

$$\text{avec } CT_{\text{input_corrige}} = CT_{\text{input}} - \log_2(FD)$$

Ici: Pour Hoxc8: $AE=1,96 \sim 2$

- Pour le ChIP: $CT_{\text{input_corrige}} = CT_{\text{input}} - 5,1$
- Pour le MeDIP: $CT_{\text{input_corrige}} = CT_{\text{input}} - 3,7$

Calculs pour les expériences de ChIP et MeDIP (2)

- Pour chaque ChIP, calculer la déviation standard du pourcentage d'input:

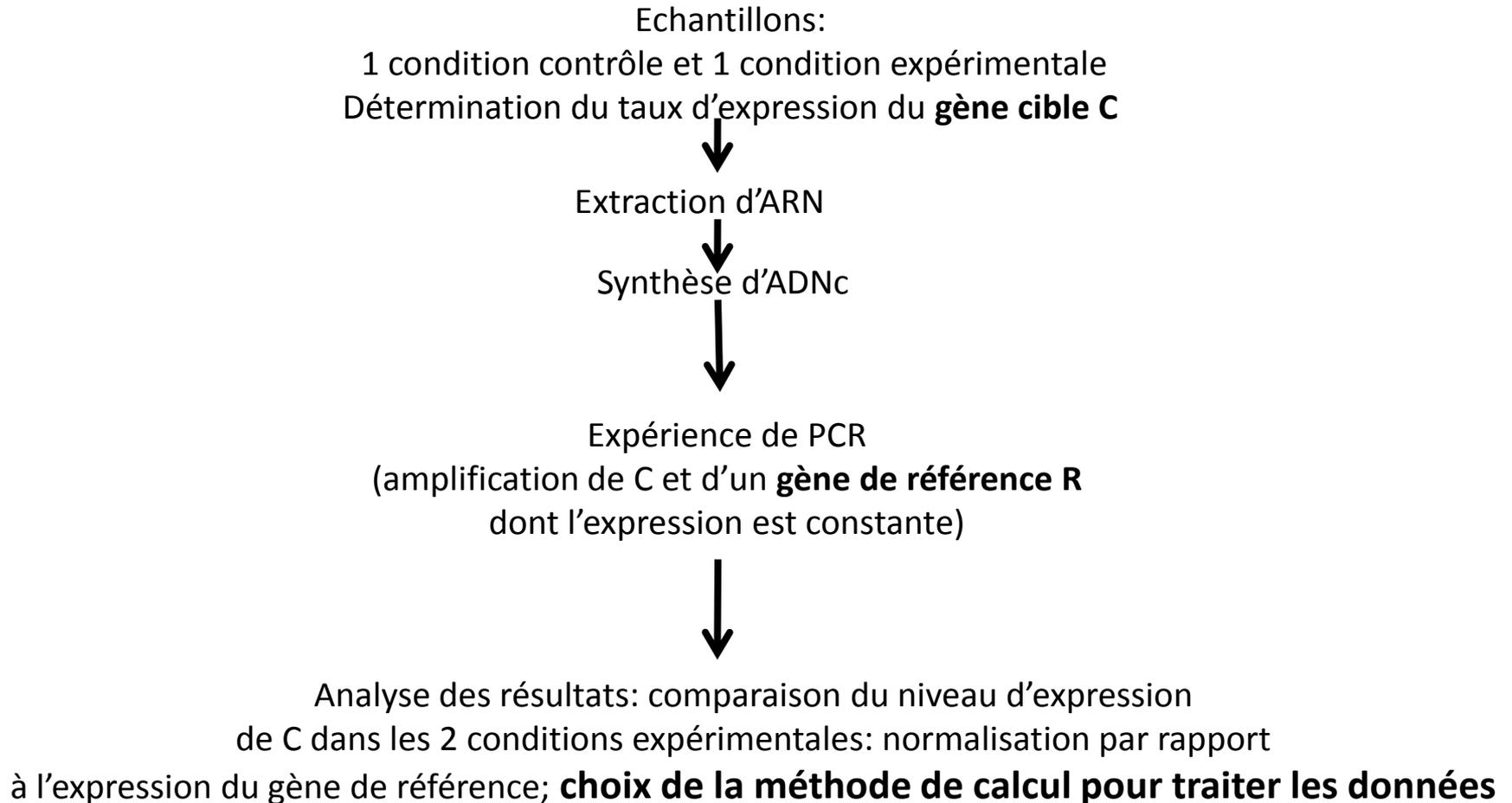
$$sd = \sqrt{(sd\#ChIP)^2 + (sd\#Input)^2}$$

- Pour n expériences répétées:

$$sd = \sqrt{(sd\#1)^2 + (sd\#2)^2 + \dots + (sd\#n)^2}$$

**Addendum: méthodes de calcul pour la Q-RTPCR en
quantification relative**

Q-RT-PCR avec quantification relative



Quelques méthodes de calcul pour la quantification relative (1)

Méthode $2^{-\Delta\Delta CT}$ (Livak *et al.*, Methods 2001, 402-408): calcul direct sur les valeurs de CT, en présumant que les efficacités du gène cible et du gène de référence sont toutes les 2 égales à 100% (AE=2)

$$\Delta C_T = C_{T \text{ gène cible}} - C_{T \text{ gène de référence}}$$

$$\Delta\Delta C_T = (\Delta C_T)_{\text{condition contrôle}} - (\Delta C_T)_{\text{condition expérimentale}}$$

Calcul de $2^{-\Delta\Delta CT}$:

Si $2^{-\Delta\Delta CT} > 1$, gène sur-exprimé dans la condition expérimentale par rapport à la condition contrôle

Si $2^{-\Delta\Delta CT} < 1$, gène sous-exprimé dans la condition expérimentale par rapport à la condition contrôle

Méthode à éviter si les efficacités ne sont pas identiques et égales à 100%:

- si efficacité cible-efficacité référence=0,03, erreur de **209%** sur le ratio d'expression après 25 cycles
- efficacité référence-efficacité cible=0,03, erreur de **47%** sur le ratio d'expression après 25 cycles

Quelques méthodes de calcul pour la quantification relative (2)

Méthode de Pfaffl (Pfaffl., NAR 2001): calcul direct sur les valeurs de CT, en tenant compte des efficacités du gène cible et du gène de référence, qui ne doivent pas forcément être identiques et égales à 100%

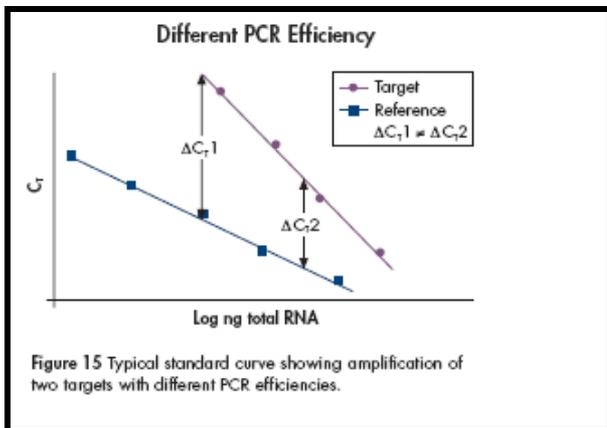
$$\Delta C_{T \text{ cible}} = C_{T \text{ condition contrôle}} - C_{T \text{ condition expérimentale}}$$

$$\Delta C_{T \text{ référence}} = C_{T \text{ condition contrôle}} - C_{T \text{ condition expérimentale}}$$

$$\text{Ratio (expérimental/contrôle)} = \frac{(AE_{\text{cible}})^{\Delta C_{T \text{ cible}}}}{(AE_{\text{référence}})^{\Delta C_{T \text{ référence}}}}$$

Exemple:

$$AE_{\text{cible}} = 1,9; AE_{\text{référence}} = 1,97$$



échantillon	CT cible	CT réf	$\Delta C_{T \text{ cible}}$	$\Delta C_{T \text{ référence}}$	Ratio
Contrôle	37,01	18,83	0	0	1
Expé 1	29,43	18,59	7,58	0,24	110,2
Expé 2	34,43	18,59	2,58	0,24	4,4

Quelques méthodes de calculs pour la quantification relative (3)

Quantification par la méthode de la courbe de calibration (ABI, user bulletin n°2): calcul des quantités X_0 à partir des C_T et de la courbe standard

-A partir des courbes de calibration, détermination de $\text{Log}(X_0)$ dans les échantillons étudiés (condition contrôle et expérimentale) pour le gène cible et le gène de référence.

-Normalisation par rapport à l'expression du gène de référence pour chaque échantillon (condition contrôle et expérimentale):

$$\text{Norm} = X_{0(\text{gène cible})} / X_{0(\text{gène référence})}$$

Ratio = Norm(expé) / Norm(contrôle):

Si $R > 1$, gène sur-exprimé dans la condition expérimentale par rapport à la condition contrôle;

Si $R < 1$, gène sous-exprimé dans la condition expérimentale par rapport à la condition contrôle

Exemple précédent:

$$AE_{\text{cible}} = 1,9; AE_{\text{référence}} = 1,97$$

échantillon	CT cible	CT réf	X_0 cible	X_0 référence	Norm	Ratio
Contrôle	37,01	18,83	$1,4 \cdot 10^{-3}$	346	$4,05 \cdot 10^{-6}$	1
Expé 1	29,43	18,59	$1,8 \cdot 10^{-1}$	417	$4,32 \cdot 10^{-4}$	106,7
Expé 2	34,43	18,59	$7,2 \cdot 10^{-3}$	417	$1,73 \cdot 10^{-5}$	4,3

Même résultat qu'avec la méthode de Pfaffl (avec la méthode de Livak: ratio (expé 1)=160; ratio expé 2=5,1)